European Journal of Human Genetics (2007), 1–7

www.nature.com/ejhg

ARTICLE

A novel sampling design to explore gene-longevity associations: the ECHA study

Francesco De Rango[1], Serena Dato[1], Dina Bellizzi[1], Giuseppina Rose[1], Erika Marzi[2], Luca Cavallone[2], Claudio Franceschi[2,6], Axel Skytthe[3], Bernard Jeune[3], Amandine Cournil[4], Jean Marie Robine[4], Jutta Gampe[5], James W Vaupel[5], Vincenzo Mari[6], Emidio Feraco[6], Giuseppe Passarino[1], Andrea Novelletto[1] and Giovanna De Benedictis*[,1]

[1]*Department of Cell Biology, University of Calabria, Rende, Italy;* [2]*C.I.G. – Interdepartmental Center 'L Galvani', University of Bologna, Italy;* [3]*Aging Research Center and Epidemiology, Institute of Public Health, University of Southern Denmark, Denmark;* [4]*Centre Regional de Lutte contre le Cancer, Montpellier, France;* [5]*Max-Planck Institute for Demographic Research, Rostock, Germany;* [6]*Italian National Research Centre on Aging (INRCA), Cosenza and Ancona, Italy*

To investigate the genetic contribution to familial similarity in longevity, we set up a novel experimental design where cousin-pairs born from siblings who were concordant or discordant for the longevity trait were analyzed. To check this design, two chromosomal regions already known to encompass longevity-related genes were examined: 6p21.3 (genes *TNFα*, *TNFβ*, *HSP70.1*) and 11p15.5 (genes *SIRT3*, *HRAS1*, *IGF2*, *INS*, *TH*). Population pools of 1.6, 2.3 and 2.0 million inhabitants were screened, respectively, in Denmark, France and Italy to identify families matching the design requirements. A total of 234 trios composed by one centenarian, his/her child and a child of his/her concordant or discordant sib were collected. By using population-specific allele frequencies, we reconstructed haplotype phase and estimated the likelihood of Identical By Descent (IBD) haplotype sharing in cousin-pairs born from concordant and discordant siblings. In addition, we analyzed haplotype transmission from centenarians to offspring, and a statistically significant Transmission Ratio Distortion (TRD) was observed for both chromosomal regions in the discordant families ($P = 0.007$ for 6p21.3 and $P = 0.015$ for 11p15.5). In concordant families, a marginally significant TRD was observed at 6p21.3 only ($P = 0.06$). Although no significant difference emerged between the two groups of cousin-pairs, our study gave new insights on the hindrances to recruiting a suitable sample to obtain significant IBD data on longevity-related chromosomal regions. This will allow to dimension future sampling campaigns to study-genetic basis of human longevity.
European Journal of Human Genetics advance online publication, 7 November 2007; doi:10.1038/sj.ejhg.5201950

Keywords: human longevity; linkage; mendelian transmission distortion

Introduction

Genetic factors involved in human aging become increasingly important[1] and specific[2] at oldest ages. Therefore, the analysis of subjects strongly selected for longevity provides a powerful tool to disentangle the complex gene network, which influences survival in the elderly.[3] Owing to the multifactorial nature of the longevity trait, sampling designs targeted at non-parametric association analyses (both case–control and sib-pair designs) are currently used to discover the genes contributing to the trait. In case–control studies, the problem is that cases (for example, centenarians) and controls (younger subjects) belong to different cohorts,

*Correspondence: Professor G De Benedictis, Department of Cell Biology, University of Calabria, Ponte Pietro Bucci, Rende 87036, Italy.
Tel: +39 0984 492 932; Fax: +39 0984 492 911;
E-mail: g.debenedictis@unical.it
Received 7 February 2007; revised 22 August 2007; accepted 28 September 2007

and this may introduce potential biases due to demographic and, to a larger extent, environmental changes occurring along the time. As compared to the case–control design, the sib-pair approach restricts the comparison within sibships, thus ensuring that subjects of each pair belong to quite close cohorts and are exposed in a comparable manner to environmental factors (at least those acting intrafamiliarly). Nevertheless, the possibility that both the sibs attained longevity more because of shared favorable environments than because of shared favorable genes should not be undervalued. A tool to check this possibility is to investigate, in addition to sib-pairs in which both sibs attained longevity (concordant pairs), sib-pairs in which a sole sib shows the longevity trait (discordant sib-pairs). In fact, both in concordant and discordant pairs, the sibs are expected to share intrafamiliar environmental factors. Therefore, the phenotypic concordance in long-living sibs should be due to both genetic and environmental factors; but if longevity is attained by a sole sib (discordant sib-pair), longevity genetic factors are expected to be restricted to the long-living sib only. Thus, different patterns of Identical By Descent (IBD) haplotype sharing are expected in concordant and discordant sib-pairs as chromosomal regions encompassing longevity genetic factors are regarded. In particular, for such regions, an over-sharing of IBD haplotypes (with respect to Mendelian expectation) should be observed in concordant sib-pairs, whereas the contrary (under-sharing) is expected in discordant sib-pairs. The use of discordant sib-pairs in linkage analysis is known in quantitative trait-locus mapping[4,5] but is never applied to explore the genetics of human longevity. Obviously, in such a case, the discordant sib is dead; however, it is possible to go down one generation and to apply IBD analysis to cousin-pairs born from sibs who are concordant or discordant for the longevity trait.

The aim of the present work was to check the feasibility and effectiveness of this novel design by using two chromosomal regions (6p21.3 and 11p15.5) encompassing genes having functional variants that are reported by literature to be associated with longevity.[6–11]

The novel design, by implying the collection of offspring of centenarians, also provided us with the opportunity to check Mendelian transmission ratio distortion (TRD)[12,13] in the studied regions. In fact, TRD, which has been defined as 'a statistically significant departure from Mendelian inheritance ratio expected regardless of the cause,'[14] has been observed in these chromosomal regions,[15,16] and this may have an effect on IBD analyses.

The study was carried out in the frame of a European project aimed at investigating genetic and environmental factors affecting aging across Europe (ECHA project; http://biologia.unical.it/echa/documents.htm). We report here the results of an exploratory genotyping effort performed on the entire DNA collection that we assembled and focused on three and five genes located in chromosomal regions 6p21.3 and 11p15.5, respectively (Table 1).

Materials and methods

Sampling

Two categories of sib-pairs were considered: concordant *versus* discordant for the longevity trait. A further requirement was that at least one child of both the proband and his/her sibling was available for sampling, in order to genotype cousin-pairs born from concordant/discordant sib-pairs. Hereafter, the proband, his/her child and the cousin are referred to as a trio. Table 2a summarizes the recruitment criteria for concordant and discordant ECHA families. Thresholds for determining eligibility for sampling were determined as follows:

Probands We set the lower age limit for longevity to the age corresponding to the surviving upper 0.05% of the population, approximately. In addition, as male/female (M/F) ratio among centenarians shows regional disparities,[29] we adopted slightly different age windows for recruiting probands in Denmark (Southern Denmark), France (Languedoc-Roussillon) and Italy (Calabria) to get an equal representation of both genders in the study. This implied a 2-year M/F difference in the Italian subsample and a 3-year M/F difference in the Danish and French subsamples.

Concordant/discordant sib-pairs The choice of age limits for the two groups was dictated by the need for obtaining a balance between a sufficient number of families and an efficient discrimination between the two groups. We then fixed thresholds based upon the survival distributions in each of the surveyed populations: if the proband's sib died at an age corresponding approximately to the range encompassing 40% of the survival in the general population, the sib-pair was considered discordant; sib's survival to an age corresponding to the upper 10% of the survival distribution determined the concordant status. This produced different cutoff ages both for men and women (due to sex-specific survival distributions) and for the three subsamples under study (Table 2a).

Identification of suitable trios As an initial step to identify sib-pairs fulfilling the above criteria, population pools of 1.6, 2.3 and 2.0 million inhabitants were screened, respectively, to completion in southern Denmark, Languedoc-Roussillon (France) and Southern Italy. As to France, as no centralized register of residents was available, general practitioners and nursing homes of the entire Languedoc-Roussillon region (2.3 millions of inhabitants) were contacted by phone about possible long-lived probands. We started from the identification of the probands according to the requirements listed in Table 2a. Then,

Table 1 Genes and polymorphisms analyzed in the present study

Gene	*OMIM*	*Chromosome*	*Polymorphism*	*Function of the gene product*	*Function of the polymorphism*
SIRT3	604481	11p15.5	VNTR intron 5[8]	Multifunctional protein localized to mitochondria upon cell stress[17] ADP-ribosyltransferase and NAD-dependent deacetylase activity	Allele-specific enhancer activity[8]
HRAS1	190020	11p15.5	VNTR intron 1[18,19]	Ras proteins bind GDP/GTP and possess intrinsic GTPase activity	Regulation on H-ras gene expression[20,21]
IGF2	147470	11p15.5	Ava II RFLP[22]	Mediation of growth hormone action and stimulation of insulin action autocrine regulator of cell proliferation	LD with the 5'INS VNTR region that regulates INS and IGF2 gene expression[23]
INS	176730	11p15.5	FokI RFLP[23]	Precursor of insulin A and B chains	LD with the 5'INS VNTR region that regulates INS and IGF2 gene expression[23]
TH	191290	11p15.5	STR intron 1[24]	Rate-limiting enzyme in the synthesis of catecholamines	Allele-specific effects on TH gene expression[25]
TNFβ and *TNFα*	153440 and 602872	6p21.3	TNFβ-a; TNFβ-c; TNFα-d; TNFα-e microsatellites[26,27]	Multifunctional proinflammatory cytokines, with effects on lipid metabolism, coagulation, insulin resistance, and endothelial function	LD within the HLA region (www.hapmap.org/cgi-pere/gbrowse/hapmap-B35/)
HSP70.1	140550	6p21.3	−110 A/C[28]	Stress-inducible molecular chaperone that controls protein folding and prevent aggregation of proteins	Lower synthesis of HSP70 protein in cell cultures from female centenarians of AA genotype[7]

LD indicates linkage disequilibrium.

Table 2 Criteria of recruitment (a) and number of samples collected (b) in Denmark, France and Italy

	Denmark		*France*		*Italy*	
a	M	F	M	F	M	F
Proband's min age	95	98	95	98	97	99
Max age of death of the proband's sib						
In concordant sib-pairs	—	—	—	—	—	—
In discordant sib-pairs	75	80	75	80	75	80
Min age of death of the proband's sib						
In concordant sib-pairs	88	94	90	93	92	94
In discordant sib-pairs	55	60	55	60	55	60
b						
Number of subjects displaying longevity phenotype						
In concordant sib-pairs	29 (96)	24 (100)	15 (98)	21 (101)	35 (98)	32 (99)
In discordant sib-pairs	22 (97)	11 (101)	11 (96)	9 (101)	22 (98)	26 (100)
Generation II						
From concordant sib-pairs	63 (65)	43 (69)	32 (70)	25 (68)	53 (71)	60 (69)
From discordant sib-pairs	40 (67)	26 (66)	21 (69)	21 (70)	45 (67)	51 (72)

M, males, F, females.
In b, the median ages of each sample category are given in parentheses. Ages are in years.

for each proband, the sampling of a complete trio was pursued; in the concordant pairs, we also sampled the proband's sib, when available.

Hindrances to sampling to completion The most common cause of exclusion from sampling turned out to be the lack of the proband's children or their unavailability for sampling. In these cases, reconstruction of the proband's pedigree was not pursued further. It is then not possible to calculate among all sibships, including a centenarian, the proportion of those fulfilling concordance or discordance.

For each trio, information on familial, social and clinical data was collected, respectively, by two questionnaires specifically assembled for probands and offspring (reported at the ECHA web site). All information was properly coded to guarantee anonymity and stored in a dedicated database.

The ECHA design and procedures were approved by the local ethical committees of the participating institutions. Written informed consent was obtained for each recruited subject. Details on aims and tools of the ECHA project, strategies for identification of the sampling targets in the three countries and procedures, and fulfillment of ethical requirements can be found at the ECHA web site (see above).

Genotyping

From each subject, 6 ml of venous blood were drawn and used for DNA preparation according to standard procedures. Figure 1 shows markers and chromosomal regions under study. Protocols used in genotyping are available on request.

Data analysis

The reconstruction of haplotype phase and the likelihood of haplotype sharing were obtained by a Maximum Likelihood algorithm included in *SimWalk2* software.[30] Recombination frequencies between markers were estimated on the basis of the physical distances showed in *MapView* (www.ncbi.nlm.nih.gov/mapview) by assuming 1 Mb = 1 cM. First, we estimated allele frequencies in each national sample by using one subject for each family (the niece/nephew of the proband). Then, we checked for possible allele frequency differences among the national groups (heterogeneity test by permutations). As the three samples had population-specific patterns of allele frequencies, both haplotype reconstruction and the following IBD analysis were carried out separately in each national sample. We calculated the likelihood of IBD haplotype sharing in each cousin-pair and counted the number of cousin-pairs having the same IBD likelihood. Possible differences between the likelihood distributions obtained in concordant and discordant cousin-pairs were checked by Mann–Whitney *U* test.

To verify whether haplotypic transmission deviated from the expected Mendelian inheritance ratio (1:1), we inferred the haplotypes transmitted (or not transmitted) from centenarians to offspring and applied a likelihood ratio test by SPSS software (Version 13.0, from SPSS Inc.).

Results

Sampling

On the whole, 851 and 1567 centenarians' families have been screened, respectively, in Denmark and Italy. As for France, 750 general practitioners and 300 nursing homes have been contacted. Our effort resulted in a total of 737 subjects from 234 families collected in the three countries (Table 2b). The discrepancy between the total number of subjects and the number of families ($234 \times 3 = 702$) is accounted for by a few concordant families with two or more sampled subjects in generation I and a few families with more than two sampled subjects in the following generation. As to generation II, the data show that the ages of the sampled subjects are very similar (Table 2b). This similarity was expected for concordant sibships. However, the fact that it is replicated in discordant sibships denotes that our samples represent subjects born at similar parental ages in the two categories of families.

Genetic analyses

All subjects were genotyped at the loci *SIRT3*, *HRAS1*, *IGF2*, *INS*, *TH*, *TNFα*, *TNFβ* and *HSP70.1*. (Table 1 and Figure 1). Out of more than 5000 genotypings, parent–offspring incompatibility was observed in less than 0.2% of the cases. Re-typing of these samples dismissed incompatibility in all cases; an analysis of Y-chromosome (father/son) and mitochondrial DNA markers (mother/offspring) confirmed parent–offspring compatibility in these cases. Furthermore, a random subset of about 20% of the samples was typed twice, producing consistent results in all cases. Therefore, the impact of genotyping errors in our study can be considered minimal because neither incompatibility nor discrepancy was observed after re-typing.

Genetic similarity in cousin-pairs

As haplotype reconstruction is highly sensitive to the background allele frequencies, we preliminarily checked if the country-specific allele frequencies allowed us to pool the three national samples. As significant heterogeneity was found among the three sample groups (Supplementary Table 1), we reconstructed the segregating haplotypes in each national sample. We then calculated the likelihood that cousin-pairs shared IBD haplotypes and checked for possible differences between the likelihood distributions of concordant and discordant cousin-pairs. Neither at the 6p21.3 region nor at the 11p15.5 region did the cousin-pairs from concordant and discordant siblings show

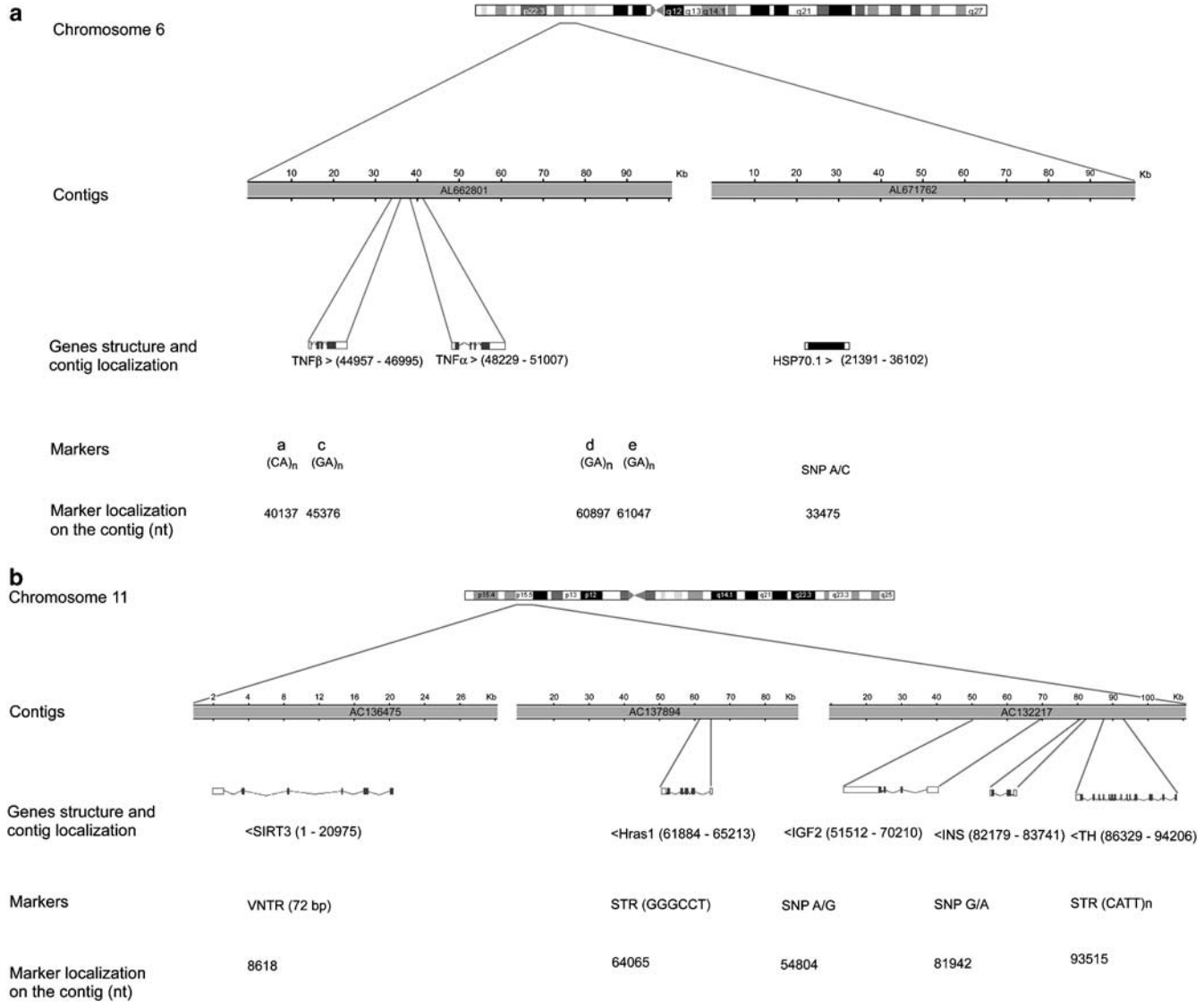

Figure 1 Schematic representation of the markers localized at 6p21.3 (**a**) and 11p15.5 (**b**) chromosomes. Chromosome structure, contig numbers, genes and markers localization are reported. For each gene, exons are represented as filled boxes. The symbols > and < indicate the forward and reverse sequences, respectively. The figures are assembled according to www.ensembl.org.

significantly different IBD patterns in any national sample (results not shown).

Transmission ratio distortion

By using country-specific allele frequencies, we first reconstructed the 6p21.3 and 11p15.5 haplotypes transmitted (or not transmitted) from each proband to his/her offspring in each national sample. Then, we analyzed Mendelian transmission in concordant and discordant families. In this case, as we were checking an extrinsic hypothesis (1:1 Mendelian transmission), we pooled the national data to increase the sample size. The results are reported in Supplementary Tables 2 and 3 (concordant and discordant families, respectively), where the rare class encompasses haplotypes present only once. In discordant families, the distribution of transmitted and not transmitted haplotypes significantly differed from the 1:1 Mendelian expectation both at 6p21.3 and 11p15.5 ($P=0.007$ and $P=0.015$, respectively), whereas in concordant families, the haplotypic transmission tended to be distorted at 6p21.3 only ($P=0.06$).

Discussion

The choice of the optimal study design to discover gene-longevity associations in humans is a real challenge.[31,32] First, because of the effects played on lifespan by social and

economic factors, longevity is a phenotype whose definition may change according to the place and year of birth of the cohort; second, true controls in candidate-gene association studies are missing; third, non-parametric linkage analyses require large sample sizes whose statistical power varies according to the dominant/recessive effect of the gene;[33] and, finally, longitudinal studies in which cohorts of individuals are followed over time are very difficult to realize. By comparing genetic similarity in cousin-pairs born from sibships of long-lived subjects (concordant sib-pairs) with cousin-pairs born from sibships in which a single subject attained longevity (discordant sib-pairs), the ECHA design aimed at controlling for potential confounding effects of stratification and genetic demographic shifts in the communities from which the subjects were sampled.

As the aim of the work was not to discover new gene-longevity associations but to explore the potentiality of a novel study design, we analyzed functional polymorphisms of genes known to be associated with longevity by case/control studies[6–11] or in strong linkage disequilibrium with functional variants affecting biological pathways involved in aging (Table 1).

The ECHA sampling design (Table 2) required a huge screening effort that produced 737 subjects from 234 families starting from a basin of about 6 million people. However, despite the strong phenotypic selection, the results of IBD analyses did not reveal significant differences between concordant and discordant families. These negative findings, which contrast with the significant gene-longevity associations found by case–control studies, are most likely explained by sample sizes inadequate to reveal small genetic effects,[33] which are further smoothed when searched in the children of the probands and their cousins (one generation away from the one selected). A further confusing effect may have been played by TRD, which has been observed in our samples for both the studied chromosomal regions. In this frame, it may be important to notice that the chromosomal regions that we studied were both known to harbor genes that play a role in longevity and turned out to have a significant TRD. Considering that a trade-off between longevity and fertility has been described[34,35] and that TRD is probably due to a low fertility of some allelic combination on specific chromosomal regions,[36] it is possible to hypothesize that the chromosomal regions involved in longevity may be more prone to TRD than others. However, further studies are needed to clarify this point. On the whole, the novel design resulted in more complexity and less efficiency in revealing gene-longevity associations in comparison with classic non-parametric approaches.[32,33,37]

In conclusion, the ECHA study failed to evidence differences of IBD haplotype sharing in cousin-pairs born from concordant and discordant sib-pairs. It is likely that a larger sample size was required. On the other hand, the strict criteria of recruitment, although applied to a potential basin of about 6 million people, have not permitted the collection of a higher number of trios matching the ECHA criteria. If our findings will be confirmed by further replication studies, new sampling strategies that focus on the recruitment of large samples of centenarians and their offspring might be planned. In any case, the fulfillment of the ECHA sampling scheme can be used to dimension future campaigns to reach a desired power and/or level of selectivity.

Acknowledgements

GDB thanks the CHRONOS project for the help given in France. This work was financially supported by EU-5 FP (ECHA project; Contract QLK6-CY-2001-00128).

References

1 Hjelmborg JV, Iachine I, Skytthe A *et al*: Genetic influence on human lifespan and longevity. *Hum Genet* 2006; **119**: 312–321.
2 Passarino G, Montesanto A, Dato S *et al*: Sex and age specificity of susceptibility genes modulating survival at old age. *Hum Hered* 2006; **62**: 213–220.
3 De Benedictis G, Franceschi C: The unusual genetics of human longevity. *Sci Aging Knowledge Environ* 2006; **28**: pe20.
4 Risch N, Zhang H: Extreme discordant sib pairs for mapping quantitative trait loci in humans. *Science* 1995; **268**: 1584–1589.
5 Forrest WF, Feingold E: Composite statistics for QTL mapping with moderately discordant sibling pairs. *Am J Hum Genet* 2000; **66**: 1642–1660. Erratum in Am J Hum Genet 2000; 66: 2020.
6 Dato S, Carotenuto L, De Benedictis G: Genes and longevity: a genetic-demographic approach reveals sex- and age-specific gene effects not shown by the case–control approach (APOE and HSP70.1 loci). *Biogerontology* 2007; **8**: 31–41.
7 Marini M, Lapalombella R, Canaider S *et al*: Heat shock response by EBV-immortalized B-lymphocytes from centenarians and control subjects: a model to study the relevance of stress response in longevity. *Exp Gerontol* 2004; **39**: 83–90.
8 Bellizzi D, Rose G, Cavalcante P *et al*: A novel VNTR enhancer within the SIRT3 gene, a human homologue of SIR2, is associated with survival at oldest ages. *Genomics* 2005; **85**: 258–263.
9 Bonafe M, Barbi C, Olivieri F *et al*: An allele of HRAS1 3′variable number of tandem repeats is a frailty allele: implication for an evolutionarily-conserved pathway involved in longevity. *Gene* 2002; **286**: 121–126.
10 De Luca M, Rose G, Bonafe M *et al*: Sex-specific longevity associations defined by Tyrosine Hydroxylase–Insulin–Insulin Growth Factor 2 haplotypes on the 11p15.5 chromosomal region. *Exp Gerontol* 2001; **36**: 1663–1671. Erratum in Exp Gerontol 2002; 37: 607–608.
11 Tan Q, Bellizzi D, Rose G *et al*: The influences on human longevity by HUMTHO1.STR polymorphism (Tyrosine Hydroxylase gene). A relative risk approach. *Mech Ageing Dev* 2002; **123**: 1403–1410.
12 Naumova AK, Greenwood CM, Morgan K: Imprinting and deviation from Mendelian transmission ratios. *Genome* 2001; **44**: 311–320.
13 Zollner S, Wen X, Hanchard NA, Herbert MA, Ober C, Pritchard JK: Evidence for extensive transmission distortion in the human genome. *Am J Hum Genet* 2004; **74**: 62–72.
14 Pardo-Manuel de Villena F, Sapienza C: Nonrandom segregation during meiosis: the unfairness of females. *Mamm Genome* 2001; **12**: 331–339.

15 Eaves IA, Bennett ST, Forster P *et al*: Transmission ratio distortion at the INS-IGF2 VNTR. *Nat Genet* 1999; **22**: 324–325.
16 Hanchard N, Rockett K, Udalova I *et al*: An investigation of transmission ratio distortion in the central region of the human MHC. *Genes Immun* 2006; **7**: 51–58.
17 Scher MB, Vaquero A, Reinberg D: SirT3 is a nuclear NAD+ dependent histone deacetylase that translocates to the mitochondria upon cellular stress. *Genes Dev* 2007; **21**: 920–928.
18 Iwahana H, Orita M, Kanazawa H, Hayashi K, Sekiya T: A new RFLP in intron 1 of the human c-Ha-ras1 gene and its close relationship with the variable tandem repeats in the region 3′ to the gene. *Oncogene* 1990; **5**: 1049–1053.
19 Tanci P, Genuardi M, Santini SA, Neri G: PCR detection of an insertion/deletion polymorphism in intron 1 of the HRAS1 locus. *Nucleic Acids Res* 1992; **20**: 1157.
20 Pethe V, Shekhar PV: Estrogen inducibility of c-Ha-ras transcription in breast cancer cells. Identification of functional estrogen-responsive transcriptional regulatory elements in exon 1/intron 1 of the c-Ha-ras gene. *J Biol Chem* 1999; **274**: 30969–30978.
21 Kotsinas A, Gorgoulis VG, Zacharatos P *et al*: Additional characterisation of a hexanucleotide polymorphic site in the first intron of human H-ras gene: comparative study of its alteration in non-small cell lung carcinomas and sporadic invasive breast carcinomas. *Cancer Genet Cytogenet* 2001; **126**: 147–154.
22 Schneid H, Girard F, Binoux M, Le Bouc Y: Ava II RFLP at the insulin-like growth factor II (IGF II) locus on chromosome 11. *Nucleic Acids Res* 1989; **17**: 466.
23 Lucassen AM, Julier C, Beressi JP *et al*: Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment of DNA spanning the insulin gene and associated VNTR. *Nat Genet* 1993; **4**: 305–310.
24 Meloni R, Albanese V, Ravassard P, Treilhou F, Mallet J: A tetranucleotide polymorphic microsatellite, located in the first intron of the tyrosine hydroxylase gene, acts as a transcription regulatory element *in vitro*. *Hum Mol Genet* 1998; **7**: 423–428.
25 Albanese V, Biguet NF, Kiefer H, Bayard E, Mallet J, Meloni R: Quantitative effects on gene silencing by allelic variation at a tetranucleotide microsatellite. *Hum Mol Genet* 2001; **10**: 1785–1792.
26 Jongeneel CV, Briant L, Udalova IA, Sevin A, Nedospasov SA, Cambon-Thomsen A: Extensive genetic polymorphism in the human tumor necrosis factor region and relation to extended HLA haplotypes. *Proc Natl Acad Sci USA* 1991; **88**: 9717–9721.
27 Udalova IA, Nedospasov SA, Webb GC, Chaplin DD, Turetskaya RL: Highly informative typing of the human TNF locus using six adjacent polymorphic markers. *Genomics* 1993; **16**: 180–186.
28 Abravaya K, Phillips B, Morimoto RI: Heat shock-induced interactions of heat shock transcription factor and the human hsp70 promoter examined by *in vivo* footprinting. *Mol Cell Biol* 1991; **11**: 586–592.
29 Robine JM, Caselli G, Rasulo D, Cournil A: Differentials in the femininity ratio among centenarians: variations between northern and southern Italy from 1870. *Popul Stud* 2006; **60**: 99–113.
30 Sobel E, Lange K: Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker sharing statistics. *Am J Hum Genet* 1996; **58**: 1323–1337.
31 Christensen K, Johnson TE, Vaupel JW: The quest for genetic determinants of human longevity: challenges and insights. *Nat Rev Genet* 2006; **7**: 436–448.
32 Tan Q, Kruse TA, Christensen K: Design and analysis in genetic studies of human ageing and longevity. *Ageing Res Rev* 2006; **5**: 371–387.
33 Tan Q, Zhao JH, Iachine I *et al*: Power of non-parametric linkage analysis in mapping genes contributing to human longevity in long-lived sib-pairs. *Genet Epidemiol* 2004; **26**: 245–253.
34 Westendorp RG, Kirkwood TB: Human longevity at the cost of reproductive success. *Nature* 1998; **396**: 743–746.
35 Lycett JE, Dunbar RI, Voland E: Longevity and the costs of reproduction in a historical human population. *Proc Biol Sci* 2000; **267**: 31–35.
36 Pardo-Manuel de Villena F, de la Casa-Esperon E, Sapienza C: Natural selection and the function of genome imprinting: beyond the silenced minority. *Trends Genet* 2000; **16**: 573–579.
37 Puca AA, Daly MJ, Brewster SJ *et al*: A genome-wide scan for linkage to human exceptional longevity identifies a locus on chromosome 4. *Proc Natl Acad Sci USA* 2001; **98**: 10505–10508.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)